Author and Year	Location	Number of Respondents	Summary of Findings		
	General Methodologic Studies (see also Bowlin (1996) under Cardiovascular Disease)				
Shea (1991)	New York	145	Description: 1989 Study of reliability for demographic and risk factors among Blacks, Whites, and Hispanics in New York City.		
			<i>Findings:</i> Prevalence estimates were highly consistent at both time periods. Individual level test-retest reliability measures were high (Kappa >0.60) for 19 demographic and risk factors, intermediate for food consumption measures (0.40-0.76), lowest for routine checkup and BP check in past two years (Kappa = 0.54 and 0.23). Consistent reliability across all three ethnic groups.		
Stein (1993)	Massachusett s	210	Description: 1992 reliability study of multiple BRFSS questions from a statewide sample and an oversample of Blacks and Hispanics.		
			Findings: No statistically significant differences for prevalence estimates for any demographic or risk factors. Individual level reliability for demographics slightly lower among minorities than among whites but still high (>0.60), but >0.70 for nearly all risk factors in all populations. Lowest individual level (but not group level) reliability for drinking and driving, 60+ drinks per month, and when last had BP checked.		
Brownson (1994)	Missouri	222	Description: 1993 study of reliability of multiple BRFSS questions from a statewide sample.		
			<i>Findings:</i> Individual level reliability estimates for demographics, chronic conditions, and risk factors were high (Kappa values 0.82-1.00). Addressed core BRFSS subjects of HTN, Diabetes, BMI, smoking status, mammography, pap tests, mammography, and digital rectal examination.		

Author and Year	Location	Number of Respondents	Summary of Findings
Stein (1995)	Massachusett s	448	Description: 1992 study of reliability of 7 demographic and 19 risk factors from a statewide sample.
			Findings: Median discordance in individual level risk estimates was 7.8% (range: 1.2%-21.8%), but was symmetrical, i.e., some estimates at second interviews were lower and some higher. Median Kappa value across all questions was 0.75 (range: 0.38-0.90). Lowest individual concordances were recency of BP examination (Kappa = 0.54), heavy alcohol consumption (Kappa = 0.31), and drinking and driving (0.30).
Stein (1996)	Massachusett s	270	Description: 1992 study of reliability of women's health questions from a statewide sample.
			Findings: This study examined the reliability of BRFSS questions on mammography, clinical breast exam, pap tests, hysterectomy, and pregnancy status. Based on Kappa statistics, concordance exceeded 80% for all areas (range: 81% for time interval since last mammogram to 97% for ever had a pap test).
Pearson (1994)	Montana	410	<i>Description:</i> 1987-1989 personal interviews of American Indians aged 18-49 living on or near 3 Montana reservations to examine characteristics of respondents with and without telephones.
			Findings: Persons without phones were more likely to report not having cholesterol check in past year (odds ratio (OR)=2.1), chronic drinking (OR = 2.1), and binge drinking (OR=1.9) even after adjusting for demographic factors. Estimates based on telephone interviews underestimate prevalence for this population.

Author and Year	<u>Location</u>	Number of Respondents	Summary of Findings
Gentry (1985)	29 states	22,236	Description: Comparison of pooled BRFSS estimates on alcohol, smoking, overweight, hypertension, and seatbelt use with national estimates from 1979 NIAAA study on alcohol, 1979 NHLBI survey, 1979 Nat'l survey of Personal Health Practices and Consequences, 1980 NHIS.  Findings: BRFSS estimates very close to national estimates for all of these risk factors (within 1-2 percentage points) except for safety belt nonuse (7 percentage point difference).
Alcohol (including	g drinking and o	driving)	
Anda (1989) [see also under Smoking]	Michigan	1,492 phone interviews 2,802 in-person interviews	<i>Description:</i> 1982 and 1983 surveys in Michigan comparing BRFSS estimates with personal interviews for binge drinking (5+ drinks on one or more occasions).
			<i>Findings:</i> Estimates of binge drinking were 0.4 percentage points lower for men and 1.0 percentage points higher for women in the BRFSS.
Smith (1990)	21 states	Median: 1,177 Range: 628-2386	Description: Correlation of state self-reported alcohol prevalence with state alcohol consumption data for 1985.
			<i>Findings:</i> Correlation coefficients between BRFSS and alcohol consumption data were generally high (0.81 for per capita alcohol consumption, 0.74 for chronic drinking [60+ drinks/month], 0.51 for drinking and driving).

Author and Year	<u>Location</u>	Number of Respondents	Summary of Findings
Robertson (1992) [see also under Safety Belts]	13 states	Not reported	Description: Comparison of 1988 BRFSS estimates for drinking and driving with percentage of fatally injured drivers from the Fatal Accident Reporting System (FARS) with any blood alcohol or illegal levels of blood alcohol.
			Findings: Low correlation between BRFSS estimates for drinking and driving with FARS data ( $r^2 = 0.20$ for illegal alcohol levels and 0.16 for any blood alcohol).
Serdula (in press)	33 states	213,842	Description: Comparison of 1987-88 BRFSS with 1989-90 alcohol questions. 1987-88 alcohol questions included beverage-specific estimates (e.g., separate questions on beer, wine, liquor), while the 1989-90 questions included only grouped questions on alcohol consumption.
			Findings: The mean number of drinks per month consumed per month among drinkers was substantially lower for the grouped alcohol beverage estimates compared to the beverage-specific estimates (37.0 vs. 29.6 for men, 17.0 vs. 13.9 for women). Prevalence of heavy drinking (>2 drinks per day for men, >1 drink per day for women) was also lower using the grouped questions (17.9% vs.10.8% for men, 15.0% vs. 10.3% for women).

Author and Year	Location	Number of Respondents	Summary of Findings
Cardiovascular D	isease (includes	obesity, cholesterol level, H	TN, diabetes, obesity)
Jackson (1992)	California	1,588 BRFSS respondents 1,512 Stanford study respondents	<i>Description:</i> Comparison of 1987 BRFSS survey with 1986 Stanford Five-City in-person study (personal interviews and physiological measures).
			Findings: BRFSS estimates were similar to Stanford study for smoking prevalence, number of cigarettes per day, self-reported cholesterol level, ever told BP high, having been prescribed medication for BP, and compliance with taking BP medications. BRFSS estimates were 30 percentage points higher for hypertensives reporting BP under control and BRFSS estimates were 1.0-2.2 points lower for body mass index.
Bowlin (1993)	New York	626	Description: 1989-90 study of BRFSS male and female respondents in upstate New York who had physiologic measurements taken for smoking, obesity, HTN, diabetes, and elevated cholesterol.
			Findings: Compared with physiologic measures, BRFSS smoking estimates were 6-10 percentage points lower, obesity estimates 10 percentage points lower, HTN prevalence (told on >1 occasion had HTN) was 14-16 percentage points lower. Diabetes prevalence was similar. Among persons reporting hypertension, BRFSS estimates were 38-50 percentage points lower for uncontrolled hypertension. BRFSS estimates were 32-33 percentage points lower than physiologic measures for elevated

cholesterol levels.

Author and Year	<u>Location</u>	Number of Respondents	Summary of Findings
Sharlin (1993)	South Dakota	418	Description: 1986 BRFSS survey estimates for hypertension (obtained via in-person interviews) compared with physiologic measures for American Indians living on a South Dakota Reservation.
			Findings: BRFSS survey definition was "ever told" had hypertension; physiologic definition was BP of 140/90. Prevalence of hypertension was 11% from the survey and 18% from measurements. Overall sensitivity was low (50%) but specificity was 92%; positive predictive value: 43%.
Giles (1995)	South Carolina	2,714	Description: 1987 survey comparing BRFSS estimates with physiologic measures from the same population for hypertension.
			Findings: Definition was "ever told" had HTN. Overall sensitivity and specificity for BRFSS was high among black and white males and females (sensitivity: 79%-82%; specificity: 88%-91%), but was sensitivity was lower among persons <40 (range: 42%-68%).

Author and Year	<u>Location</u>	Number of Respondents	Summary of Findings
Bowlin (1996)	New York	626	Description: Reliability and validity study using 1989-90 BRFSS male and female respondents in upstate New York for CVD risk factors (see previous Bowlin citation).  Findings: Reliability of estimates was high (Kappa >0.60) for ever told BP high, current smoking, number of cigarettes/day, diabetes, ever had cholesterol checked, trying to lose weight, weight, height, and time since last checkup. Reliability was lower (Kappa <0.50) for hypertension under control and numeric values for blood pressure. Sensitivity and specificity for BRFSS estimates were high for smoking, obesity, and diabetes (range: 75% to 99%) but much lower for hypertension and elevated cholesterol level (range: 47% to 82%).

Author and Year	Location	Number of Respondents	Summary of Findings
Mammography			
Degnan (1992)	North Carolina	973	Description: Compared BRFSS question on receipt of a mammogram in the past year with estimates based on records from institutions doing mammography in New Hanover county, NC in 1987 and 1989 (note: this was a special study).  Findings: Telephone estimates were higher than facilities' records by 15 percentage points in 1987 and by 19 percentage points in 1989. Among 164 telephone respondents whose mammograms were confirmed, telescoping (remembering that mammography
			occurred more recently than it did) accounted for much of this.
Nutrition			
Serdula (1993)	Arizona Georgia Illinois Wisconsin	1,052	Description: Study conducted in four locations between 1985-1991 compared 6 fruit and vegetable items on BRFSS with data from food frequency questionnaires and diet recalls/multiple diet records.
			<i>Findings:</i> Spearman correlation coefficients between BRFSS questions and food frequency questionnaires were generally high (range: 0.47-0.57), but slightly lower when compared to diet recalls/records (range: 0.29-0.54).
Smith-Warner (1997)	Minnesota	101	Description: 1991-1994 study was a three-way comparison between BRFSS fruit/vegetable questions, food frequency surveys, and diet records.
			Findings: BRFSS underestimated average daily fruit/vegetable consumption compared to the other methods (average for BRFSS = 3.8, food frequency questionnaire = 6.5, diet recall = 6.3). Correlation between BRFSS and food frequency questionnaire was 0.63 and between BRFSS and diet records was 0.56.

Author and Year	Location	Number of Respondents	Summary of Findings
Shea (1991)	New York	145	Description: Unpublished data in this article reports on correlation between BRFSS dietary fat/cholesterol consumption and Willett food frequency questionnaire.
			<i>Findings:</i> Correlation coefficients were very high between the two methods, ranging from 0.82 to 0.97.
Coates (1995)	Arizona Georgia Illinois Wisconsin	1,052	Description: Study conducted in four locations between 1985-1991 compared dietary fat items on BRFSS with data from food frequency questionnaires and diet recalls/multiple diet records.
			<i>Findings:</i> Spearman correlation coefficients between BRFSS questions with food frequency questionnaires and diet recalls were moderate (range: 0.33-0.60). Correlations were lower for the percentage of energy intake derived from dietary fat (range: 0.26-0.42).
Serdula (1995)	21 states	231,852	Description: Study examined the effects of changing the order of questions on self-reported weight and trying to lose weight using 1985-1992 BRFSS data.
			Findings: The prevalence of women reporting that they were trying to lose weight increased from 41% to 48%, while for men, it increased from 26% to 29% when the self-reported weight question proceeded the "trying to lose weight" question.

Author and Year	Location	Number of Respondents	Summary of Findings
<b>Quality of Life</b>			
Verbrugge (1995)	49 states	20,029	Description: Compared self-rated health with persons having one or more disability days in the past 30 days.  Findings: Respondents with lower levels of self-reported health were more likely to report having had 1+ disability days, even after demographic and other characteristics were controlled for. (Odds ratios ranged from 1.3 for those with "very good" health to 4.9 for those reporting "poor" health.
Hennessey (1994)	6 states	2,961	<i>Description:</i> Compared self-reported health status with the number of poor physical or mental health days using 1993 data from 6 states.
			Findings: Persons with 1+ poor physical or mental health days were much more likely to have lower levels of self-reported health
Newschaeffer (1997)	Missouri	401	<i>Description:</i> Compared individual level data for BRFSS questions with the Medical Outcomes Short Form 36 (SF-36) in 1996.
			<i>Findings:</i> BRFSS questions, especially those on self-reported health status and poor physical health days, correlate very strongly with SF-36 score.
Beatty (in press)	Maryland	18	Description: Cognitive testing of the 4 BRFSS core instrument Quality of Life questions.
			Findings: Nearly all subjects had difficulty providing a "number of days" poor physical or mental health response. Generally, they described how they felt without quantification. It is likely that responses to these questions is highly dependent upon prompting by interviewers.

Author and Year	Location	Number of Respondents	Summary of Findings
Safety Belt Use			
CDC (1988)	15 states	Not reported	Description: Comparison between self-reported BRFSS data with state observation data for 15 states in 1987.
			Findings: Median BRFSS estimates were 8 percentage points higher than observation study estimates when safety belt use was defined as "always use" and 21.5 percentage points higher when defined as "always or nearly always use" seat belts.
Robertson (1992) [see also under Alcohol]	13 states	Not reported	Description: Comparison of 1988 BRFSS estimates for safety belt use with observational studies conducted in cities within 13 states.
Alcohor			Findings: Median BRFSS estimates were 21.5 percentage points higher when defined as "always or nearly always use" seat belts.
Nelson (1996)	49 states in 1992	Median sample size: 2008	Description: Comparison between 1992 and 1993 BRFSS estimates with state observational surveys of safety belt use.
	50 states in 1993		<i>Findings:</i> Median BRFSS estimates (based on "always use" definition) were 5% higher in 1992 and 2% higher in 1993 than observational estimates of safety belt use. Correlation between self-reported and observational data were high ( $r^2 = 0.59$ in 1992 and 0.67 in 1993)
Smoking			
Anda (1989) [see also under Alcohol]	Michigan	1,492 phone interviews 2,802 in-person interviews	Description: 1982 and 1983 surveys in Michigan comparing personal interviews with BRFSS estimates on current smoking.
. neonor		interviews	Findings: Estimates of current smoking prevalence were 2.1 percentage points lower for men and 1.3 percentage points lower for women in the BRFSS.

Author and Year	<u>Location</u>	Number of Respondents	Summary of Findings
Arday (in press)	22 states in 1985 40 states in 1989	1985 BRFSS: 25,192 CPS: 51,368	Description: Comparison of BRFSS state estimates for current smoking with estimates from the Census Bureau's Current Population Surveys (these surveys are conducted by phone and inperson).
	49 states in 1992-93	1989 BRFSS: 66,719 CPS: 97,534	Findings: Overall BRFSS median estimates were slightly lower than CPS estimates (range across years: 0.7 to 2.0 percentage points lower), and this was greater for males than females. CPS and BRFSS estimates were very similar for Blacks (range for
		BRFSS: 195,227 CPS: 269,750	median estimates: 1.1 to 2.9 percentage points lower) and Hispanics (range for median estimates: 0.5 to 1.4 percentage points higher).

## BRFSS-SPECIFIC RELIABILITY AND VALIDITY REFERENCES

Anda RF, Williamson DF, Dodson D, Remington PL: Telephone versus in-person reporting of smoking and alcohol use: A comparison of two statewide surveys. Am J Health Promotion 1989;4:32-36.

Arday DR, Tomar SL, Nelson DE, Merritt RK, Schooley MW, Mowery P. State smoking prevalence estimates: a comparison between the Behavioral Risk Factor Surveillance System and Current Population Surveys. Am J Public Health (in press).

Beatty P, Schechter S, Whitaker K. Evaluating subjective questions: cognitive and methodologic investigations. Proceeding of the 1996 American Statistical Association, Section on Survey Research Methods (in press).

Bowlin SJ, Morrill BD, Nafziger AN, Jenkins PL, Lewis C, Pearson TA. Validity of cardiovascular disease risk factors assessed by telephone survey: The Behavioral Risk Factor Survey. J Clin Epidemiol 1993;46:561-571.

Bowlin SJ, Morrill BD, Nafziger AN, Lewis C, Pearson TA. Reliability and changes in validity of self-reported cardiovascular disease risk factors using dual response: the Behavioral Risk Factor Survey. J Clin Epidemiol 1996;49:511-517.

Brownson RC, Jackson-Thompson J, Wilkerson JC, Kiani F. Reliability of information on chronic disease risk factors collected in the Missouri Behavioral Risk Factor Surveillance System. Epidemiology 1994;5:545-549.

Centers for Disease Control. Comparison of observed and self-reported seat belt use rates-United States. MMWR 1988;37:549-51.

Coates RJ, Serdula MK, Byers T, Mokdad A, Jewell S, Leonard SB, Ritenbaugh C, Newcomb P, Mares-Perlman J, Chavez N, and Block G. A brief, telephone-administered food frequency questionnaire can be useful for surveillance of dietary fat intakes. J Nutr 1995;142:1473-1483.

Degnan D, Harris R, Ranney J, et al. Measuring the use of mammography: Two methods compared. Am J Public Health 1992;82:1386-88.

Gentry EM, Kalsbeek WD, Hogelin GC, Jones JT, Gaines KL, Forman MR, Marks JS, Trowbridge FL. The Behavioral Risk Factor Surveys: II. Design, methods, and estimates from combined state data. Am J Prev Med 1985;1(6):9-14.

Giles WH, Croft JB, Keenan NL, Lane MJ, Wheeler FC. The validity of self-reported hypertention and correlates of hypertension awareness among blacks and whites within the stroke belt. Am J Prev Med 1995; 11(3):163-169

Hennessy CH, Moriarty DG, Zack MM, Scherr PA, Brackbill R. Measuring health-related quality of life for public health surveillance. Public Health Reports 1994;109(5):665-672.

Jackson C, Jatulis DE, Fortmann SP. The Behavioral Risk Factor Survey and the Stanford Five-City Project Survey: A comparison of cardiovascular risk behavior estimates. Am J Public Health 1992;82:412-416.

Nelson DE. Validity of self reported data on injury prevention behavior: Lessons from oservational and self reported surveys of safety belt use in the U.S. Injury Prevention 1996; 2(1):67-69.

Newschaeffer CJ. Validation of BRFSS health-related quality of life questions in a statewide sample. Atlanta, GA: CDC Seminar on Quality of Life measures, May 2, 1997.

Pearson D, Cheadle A, Wagner E, Tonsberg R, Psaty BM. Differences in sociodemographic, health status, and lifestyle characteristics among American Indians by telephone coverage. Prev Med 1994;23:461-464.

Robertson LS. The Validity of self-reported behavioral risk factors: seatbelt and alcohol use. The Journal of Trauma 1992, 32(1):58-59.

Serdula M, Coates R, Byers T, et al. Evaluation of a brief telephone questionnaire to estimate fruit and vegetable consumption in diverse study populations. Epidemiology 1993;4:455-463.

Serdula MK, Mokdad AH, Pamuk ER, Williamson DF, Byers T. Effects of question order on estimates of the prevalence of attempted weight loss. Am J Epidemiol 1995;142:64-67.

Serdula MK, Mokdad AH, Byers T, Siegel PZ. Assessing alcohol consumption: beverage-specific vs. grouped-beverage questions. J Studies Alcohol (in press).

Sharlin KS, Heath GW, Ford ES, Welty TK. Hypertension and blood pressure awareness among American Indians of the Northern Plains. Ethnicity Dis 1993;3:337-343.

Shea S, Stein AD, Lantigua R, Basch CE. Reliability of the Behavioral Risk Factor Survey in a triethnic population. Am J Epidemiol 1991;133:489-500.

Shea S, Stein AD, Basch CE, Lantigua R, Maylahn C, Strogatz DS, Novick L. Independent associations of educational attainment and ethnicity with Behavioral Risk Factors for cardiovascular disease. Am J Epidemiol 1991;134:567-582.

Smith PF, Remington PL, Williamson DF, Anda RF. A comparison of alcohol sales data with survey data on self-reported alcohol use in 21 states. Am J Public Health 1990;80:309-312.

Smith-Warner SA, Elmer PJ, Fosdick L, Tharp TM, Randall B. Reliability and comparability of three dietary assessment methods for estimating fruit and vegetable intakes. Epidemiol 1997;8:196-201.

Stein AD, Lederman RI, Shea S. The Behavioral Risk Factor Surveillance System questionnaire: Its reliability in a statewide sample. Am J Public Health 1993;83:1768-72.

Stein AD, Courval JM, Lederman RI, Shea S. Reproducibility of responses to telephone interviews: demographic predictors of discordance to risk factor status. Am J Epidemiol 1995;141:1097-1106.

Stein AD, Lederman RI, Shea S. Reproducibility of the women's health module of the Behavioral Risk Factor Surveillance System questionnaire. Ann Epidemiol 1996;6:47-52.

Verbrugee LM, Merrill SS, Liu X. Measuring disability with parsimony. Hyattsville, MD: Proceedings of the 25th Public Health Conference on Records and Statistics, 1995.